Knowledge Exchange: Roundtable Discussions

MODERATORS:
Anna Jester
eJournalPress
Rockville, Maryland

SPEAKER:
Angie Hunter
American Chemical Society
Washington, DC

Gregory Schwarz
American Astronomical Society
Washington, DC

REPORTER:
Taryn Dollings
American Society of Civil
  Engineers
Reston, Virginia

Publishers are increasingly striving to help authors make accurate and accessible data available not only to reviewers but also to readers. Many are considering workflow changes and new partnerships to help achieve this goal. To this end, this roundtable discussion brought together two experienced professionals to share their perspectives on validating and publishing data.

Angie Hunter, Development Editor at the American Chemical Society, delivered a presentation on "Data Review and Initiatives at *Organic Letters*," in which she reviewed the journal's data analysis program. *Organic Letters* encourages, but does not require, its authors to submit raw data with their manuscripts. In response to often missing and incorrect data (and less often, manipulated data), Hunter and her team have developed a process for the editorial office to analyze data for accuracy and quality. Most of the work takes place once a revision is requested and then submitted. Data is reviewed for consistency with the submitted manuscript, and experimental procedures are checked for accuracy. The
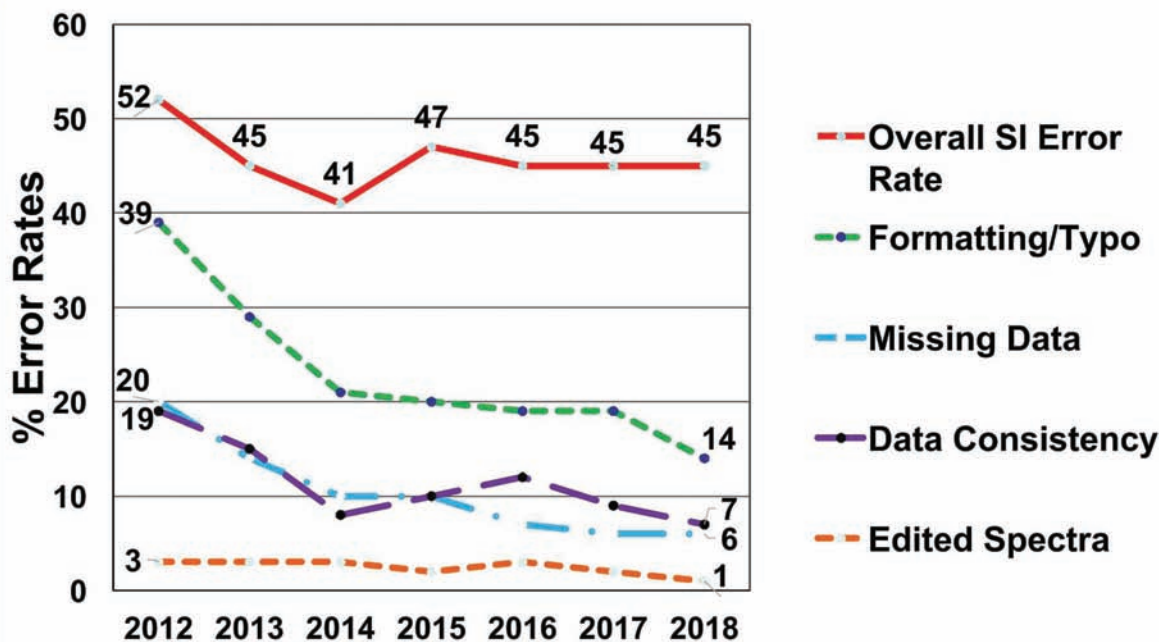
CONTINUED

editorial office also reviews images for accuracy, legibility, and possible falsification.

Hunter emphasized that although edited data or spectra comprise the smallest percentage of data analysis problems, they are often the most time-consuming to deal with. She also emphasized that the response from authors who have submitted edited data or spectra is mostly positive; authors are generally glad to be made aware of inaccuracies and have the opportunity to correct them.

In addition to checking manuscript data, Hunter's office has developed checklists and standards for authors to follow when submitted to *Organic Letters*. Author education efforts such as these have helped reduce the editorial office's data analysis workload. Figure 1 shows how *Organic Letters* has seen a reduction in various types of data-related errors from 2012 to 2018.

Gregory Schwarz, Data Editor at the American Astronomical Society (AAS), gave a presentation titled "Data behind the Figure and Interactive Figures in the American Astronomical Society Journal Articles", focusing on two of many data products that AAS produces. The Data behind the Figure (DbF) program ensures that the data used in an article figure is available in common formats for preservation and reuse, ultimately increasing article citations. In AAS's process, one of two data editors will convert the data provided into an acceptable format for publication, to be verified by the authors. AAS authors may also submit interactive figures to provide further clarity to readers. Authors may provide data, JavaScript, and HTML components that conform to AAS's publishing platform, and these figures are made available both in the published article and in the separate Astronomy Image Explorer on AAS's platform.

Schwarz emphasized the importance of flexibility and innovation as data formats and standards change over time, citing several examples from his tenure at AAS. He also encouraged publishers to adopt best standards for data formatting and sharing and partner with appropriate third parties to provide greater convenience for authors and staff. His final advice was that publishers set priorities that allow them to realistically manage the workload of formatting and verifying data, and to not let the fear of failure prevent them from pursuing better data availability and quality.

Hunter and Schwarz both have experience developing robust data analysis and curation programs and emphasized the value of having dedicated staff for data analysis and curation; subject area expertise is also a significant boon. They also stressed the importance of clear and consistent communication with authors and internal editorial teams.

Hunter and Schwarz's discussions illustrated the variety of workflow options available to publishers interested in more intensively curating data for publication. Both made convincing arguments for the value of such work while also realistically outlining its challenges. Data transparency and sharing are becoming more of an expectation every year; publishers will need to find models for data curation that work for their authors and their budgets.